Available online at www.sciencedirect.com

SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 382 (2004) 247–269

www.elsevier.com/locate/laa

Burnside's theorem: irreducible pairs of transformations

W.E. Longstaff

School of Mathematics and Statistics, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

Received 6 May 2003; accepted 29 December 2003

Submitted by R.A. Brualdi

Abstract

By Burnside's theorem, if the linear transformations A and B, acting on a finite-dimensional complex vector space $\mathscr{H}$, have no common nontrivial invariant subspaces, the words in A and B span $\mathscr{B}(\mathscr{H})$. Call the *minimum spanning length* of the pair $\{A, B\}$ the smallest positive integer l with the property that words in A and B of length at most l span $\mathscr{B}(\mathscr{H})$. Let $\operatorname{msl}(A, B)$ denote the minimum spanning length. If $\dim \mathscr{H} = 2$, $\operatorname{msl}(A, B) = 2$ and if $\dim \mathscr{H} = 3$, $\operatorname{msl}(A, B) = 3$ or 4. If $\dim \mathscr{H} \geqslant 4$, $\operatorname{msl}(A, B) \leqslant n^2 - 3$. If $\dim \mathscr{H} = n \geqslant 2$ then (i) $\operatorname{msl}(A, B) = 2n - 2$ if $\{A, B, AB, BA\}$ is linearly dependent, (ii) if B is unicellular, then $\operatorname{msl}(A, B) \leqslant 2n - 2$, where the inequality is sharp, and it can happen that $\operatorname{msl}(A, B) = n$.

AMS classification: 15A30; 47L05

Keywords: Irreducible; Span; Words

1. Introduction and preliminaries

By Burnside's theorem (see [3, Theorem 1.2.2]), if $\mathscr{T}$ is an irreducible set of (linear) transformations acting on a finite-dimensional complex vector space $\mathscr{H}$, the algebra generated by $\mathscr{T}$ is $\mathscr{B}(\mathscr{H})$, the set of all transformations on $\mathscr{H}$. Here, and in what follows, by *irreducible* we mean *having no common nontrivial invariant*

E-mail address: longstaf@maths.uwa.edu.au (W.E. Longstaff).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2003.12.043

subspace. The algebra generated by $\mathscr{T}$ is the (linear) span of the set of words with factors belonging to $\mathscr{T}$. It seems interesting to consider, for any given irreducible set $\mathscr{T}$, the smallest positive integer l with the property that the words (with factors) in $\mathscr{T}$ of length at most l span $\mathscr{B}(\mathscr{H})$. Here, by the length of a word in $\{A, B, C, \ldots\}$ we mean the number of factors in the word, counting multiplicities. For example, the word $B^2ABAC^3A^4$ has length 12. This positive integer l will be called the *minimum spanning length* of $\mathscr{T}$, denoted $\mathrm{msl}(\mathscr{T})$. An upper bound for $\mathrm{msl}(\mathscr{T})$ is not too difficult to find as the following proposition, and its corollary, show. These are due to Radjavi who has kindly permitted their inclusion here.

Proposition 1 [Radjavi]. *Let $\mathscr{T}$ be a set of transformations on a finite-dimensional complex vector space $\mathscr{H}$. Let $\mathscr{A}$ be the algebra generated by $\mathscr{T}$ and let the dimension of $\mathscr{A}$ be a and let the dimension of the span of $\mathscr{T}$ be s. Then $\mathscr{A}$ is spanned by words* (*with factors*) *in $\mathscr{T}$ of length at most $a - s + 1$.*

Proof. For every $p \geqslant 1$ let $\mathscr{W}_p$ be the span of the words in $\mathscr{T}$ of length at most p and let $s_p = \dim \mathscr{W}_p$. Then

$$\mathscr{W}_1 \subseteq \mathscr{W}_2 \subseteq \mathscr{W}_3 \subseteq \cdots \subseteq \mathscr{A},$$

so $s_1 \leqslant s_2 \leqslant s_3 \leqslant \cdots \leqslant a$. Note that if $\mathscr{W}_{q+1} = \mathscr{W}_q$, then $\mathscr{W}_{q+2} = \mathscr{W}_{q+1}$, so $\mathscr{W}_{q+2} = \mathscr{W}_q$ and $\mathscr{W}_r = \mathscr{W}_q$ for all $r \geqslant q$. There must exist a p such that $s_{p+1} = s_p$. Let the smallest such p be p_0. Then $\mathscr{W}_r = \mathscr{W}_{p_0}$ for all $r \geqslant p_0$. Since $\mathscr{A} = \bigcup_{r=1}^{\infty} \mathscr{W}_r$ it follows that $\mathscr{A} = \mathscr{W}_{p_0}$ and $a = s_{p_0}$. Now $s_1 = s$, $s_2 \geqslant s + 1$, $s_3 \geqslant s + 2, \ldots, s_{p_0} = a \geqslant s + p_0 - 1$. Hence $p_0 \leqslant a - s + 1$ and this completes the proof. $\square$

Corollary 1. *Let $\mathscr{T}$ be an irreducible set of transformations on a finite-dimensional complex vector space $\mathscr{H}$. If $\dim \mathscr{H} = n$, then $\mathscr{B}(\mathscr{H})$ is spanned by words in $\mathscr{T}$ of length at most $n^2 - s + 1$, where s is the dimension of the span of $\mathscr{T}$. Consequently* $\mathrm{msl}(\mathscr{T}) \leqslant n^2 - s + 1$.

In this paper we are primarily concerned with the minimum spanning lengths of irreducible pairs $\{A, B\}$ of transformations. Note that, for every integer $k \geqslant 1$, there are 2^k words in A and B of length k, and $2^{k+1} - 2$ words of length at most k. By the preceding corollary we have, for such pairs, $\mathrm{msl}(A, B) \leqslant n^2 - 1$ (since $\{A, B\}$ is clearly linearly independent), where $n = \dim \mathscr{H}$. We show that, not surprisingly, $\mathrm{msl}(A, B) = 2$ if $\dim \mathscr{H} = 2$, and $\mathrm{msl}(A, B) = 3$ or 4 if $\dim \mathscr{H} = 3$. Also, if $\dim \mathscr{H} = n \geqslant 2$ we show that (i) $\mathrm{msl}(A, B) = 2n - 2$ if $\{A, B, AB, BA\}$ is linearly dependent and that (ii) if B is unicellular, then $\mathrm{msl}(A, B) \leqslant 2n - 2$, where the inequality is sharp, and that it can happen that $\mathrm{msl}(A, B) = n$.

Throughout what follows $\mathscr{H}$ will denote a complex finite-dimensional vector space. If $\mathscr{H}$ is a Hilbert space the inner-product on $\mathscr{H}$ will be denoted by $(\cdot|\cdot)$.

A subspace of $\mathscr{H}$ is nontrivial if it is nonzero and different from $\mathscr{H}$. If $\{e, f, g, \ldots\}$ is a set of vectors of $\mathscr{H}$ we let $\langle e, f, g, \ldots\rangle$ denote the (linear) span of $\{e, f, g, \ldots\}$. If $\mathscr{H} = \mathbb{C}^n$ we identify $\mathscr{B}(\mathscr{H})$ with the set $M_n(\mathbb{C})$ of $n \times n$ complex matrices, more precisely, we identify a transformation T on $\mathbb{C}^n$ with its matrix relative to the standard basis for $\mathbb{C}^n$. The standard basis for $\mathbb{C}^n$ will, as usual, be denoted by $\{e_1, e_2, \ldots, e_n\}$. For each $n \geqslant 2$ and for $1 \leqslant i, j \leqslant n$, the $n \times n$ elementary matrix $E_{i,j}$ is the matrix having i, j-entry equal to one with all other entries zero. Also, the $n \times n$ strictly upper triangular elementary Jordan matrix is the matrix having ones on the first superdiagonal and zeros elsewhere, that is, each $i, i+1$-entry equals one, $1 \leqslant i \leqslant n-1$, and all other entries are zero. For $T \in \mathscr{B}(\mathscr{H})$, $\mathscr{R}(T)$ denotes the range of T. The descriptive results that we use concerning the lattice of invariant subspaces of a linear transformation can be found in [1].

2. Low dimensional cases

In this section we consider irreducible pairs of transformations on spaces of dimension 2 or 3.

Lemma 1. *Let* $\dim \mathscr{H} \geqslant 2$ *and let* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible. Then*

(a) $\{A, B, AB\}$ *is linearly independent*,
(b) *if* $B^2 \in \langle A, B, AB\rangle$, *then* B^2 *is a scalar multiple of* B. *If* $A^2 \in \langle A, B, AB\rangle$, *then* A^2 *is a scalar multiple of* A.

Proof. (a) Clearly $\{A, B\}$ is linearly independent. Suppose that $AB = \alpha A + \beta B$ for some scalars α and β. Then $(A - \beta I)(B - \alpha I) = \alpha\beta I$. We must have $\alpha\beta = 0$, otherwise $A - \beta I$ and $B - \alpha I$ commute. Thus $(A - \beta I)(B - \alpha I) = 0$ so $\mathscr{R}(B - \alpha I)$ is invariant under A and B. This contradicts irreducibility.

(b) Let $B^2 = \alpha A + \beta B + \gamma AB$ for some scalars α, β, γ.

Suppose that B is not nilpotent. Let $Bf = \lambda f$ with $\lambda \neq 0$ and $f \neq 0$. Then $\lambda^2 f = \alpha Af + \beta\lambda f + \gamma\lambda Af$. Since Af cannot be a multiple of f, $\alpha = -\gamma\lambda$ and $\lambda = \beta$. Hence $B^2 = -\gamma\lambda A + \lambda B + \gamma AB$ so $B(B - \lambda I) = \gamma A(B - \lambda I)$. Now $\mathscr{R}(B - \lambda I)$ is nonzero and invariant under B so $B(B - \lambda I)g = \delta(B - \lambda I)g$ for some vector g satisfying $(B - \lambda I)g \neq 0$ and some scalar δ. Then $\delta(B - \lambda I)g = \gamma A(B - \lambda I)g$ so $\gamma = 0$, by irreducibility. Thus $B^2 = \beta B$.

Suppose that B is nilpotent. Then $\alpha = 0$ (since if $Bh = 0$ then $\alpha Ah = 0$) so $B^2 = \beta B + \gamma AB$. Let $p \geqslant 2$ be the index of nilpotency of B. Then $0 = B^p = \beta B^{p-1} + \gamma AB^{p-1}$, so $\beta I + \gamma A$ leaves $\mathscr{R}(B^{p-1})$ invariant. Since $\mathscr{R}(B^{p-1})$ is nontrivial, $\gamma = 0$ and so $B^2 = \beta B$.

Finally, let $A^2 \in \langle A, B, AB\rangle$. Then, taking adjoints with respect to any chosen inner-product on $\mathscr{H}$, we have $(A^*)^2 \in \langle B^*, A^*, B^*A^*\rangle$ so $(A^*)^2$ is a scalar multiple of A^* by what has just been proved. $\square$

Proposition 2. *Let* $\dim \mathscr{H} = 2$ *and let* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible. Then* $\mathscr{B}(\mathscr{H})$ *is spanned by words in* A *and* B *of length at most* 2. *Consequently* $\mathrm{msl}(A, B) = 2$.

Proof. By Lemma 1, $\{A, B, AB\}$ is linearly independent. Suppose that the words of length at most 2 do not span $\mathscr{B}(\mathscr{H})$. Then A^2 and B^2 both belong to $\langle A, B, AB\rangle$ and so, by Lemma 1, $A^2 = \alpha A$ and $B^2 = \beta B$ for some scalars α and β. Also, $BA \in \langle A, B, AB\rangle$ so letting $Be = 0$ with $e \neq 0$ we have $BAe \in \langle Ae, Be, ABe\rangle = \langle Ae\rangle$. But $Ae \neq 0$ and $A(Ae) = \alpha Ae$. This contradicts irreducibility. $\square$

Next we show, in Theorem 1 below, that $\mathrm{msl}(A, B) \leqslant 4$ for every irreducible pair $\{A, B\}$ on three-dimensional space. To do this we need the following four lemmas. We prove these lemmas by showing that, in each case, the span of the words of length at most 4 in A and B contains all of the 3×3 elementary matrices. To do this we use the fact that, if the span of a set $\mathscr{T}$ of $n \times n$ matrices contains each elementary matrix in the set $\{E_{i,j} : (i, j) \in \mathscr{S}\}$ (where $\mathscr{S}$ is a subset of $\{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$) and it contains a matrix the positions of whose nonzero entries form a subset of $\mathscr{S}$ with one exception, say (u, v), then $E_{u,v}$ also belongs to the span of $\mathscr{T}$. The following lemma is a special case of Theorem 2. We include a proof here because it is more direct, and because it indicates how the general, $n \times n$ case, is proved.

Lemma 2. *Let* $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. *For any* 3×3 *matrix* A *such that* $\{A, B\}$ *is irreducible,* $M_3(\mathbb{C})$ *is spanned by words in* A *and* B *of length at most* 4.

Proof. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ and let $\mathscr{W}$ be the span of the set of words in A and B of length at most 4. Note that

$$B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & a & b \\ 0 & d & e \\ 0 & g & h \end{pmatrix}, \quad BA = \begin{pmatrix} d & e & f \\ g & h & i \\ 0 & 0 & 0 \end{pmatrix},$$

$$BAB = \begin{pmatrix} 0 & d & e \\ 0 & g & h \\ 0 & 0 & 0 \end{pmatrix}, \quad B^2A = \begin{pmatrix} g & h & i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad AB^2 = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & d \\ 0 & 0 & g \end{pmatrix},$$

$$B^2AB = \begin{pmatrix} 0 & g & h \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad BAB^2 = \begin{pmatrix} 0 & 0 & d \\ 0 & 0 & g \\ 0 & 0 & 0 \end{pmatrix}.$$

The nontrivial invariant subspaces of B are $\langle e_1\rangle$ and $\langle e_1, e_2\rangle$. By irreducibility at least one of $\{d, g\}$ is nonzero and at least one of $\{g, h\}$ is nonzero.

Case 1. $g \neq 0$. Since $B^2 = E_{1,3}$, $E_{1,3}$ belongs to $\mathscr{W}$. Then, since $B^2AB = gE_{1,2} + hE_{1,3}$, $E_{1,2} \in \mathscr{W}$, and it follows from the form of B^2A that $E_{1,1} \in \mathscr{W}$. Similarly, the forms of BAB^2, BAB and BA show that $E_{2,3}$, $E_{2,2}$ and $E_{2,1}$ belong to $\mathscr{W}$. Finally, AB^2, AB, A show that $E_{3,3}$, $E_{3,2}$, $E_{3,1} \in \mathscr{W}$.

Case 2. $g = 0$. Then d and h are both nonzero, by irreducibility. Note that

$$A^2 = \begin{pmatrix} * & * & * \\ * & * & * \\ dh & * & * \end{pmatrix}, \quad BA^2 = \begin{pmatrix} * & * & * \\ dh & * & * \\ 0 & 0 & 0 \end{pmatrix},$$

$$B^2A^2 = \begin{pmatrix} dh & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^2B = \begin{pmatrix} 0 & * & * \\ 0 & * & * \\ 0 & dh & * \end{pmatrix},$$

$$A^2B^2 = \begin{pmatrix} 0 & 0 & * \\ 0 & 0 & * \\ 0 & 0 & dh \end{pmatrix},$$

where '$*$' denotes a scalar, possibly zero. Arguing in the same way as in Case 1, B^2, B^2A, B^2A^2 show that $E_{1,3}$, $E_{1,2}$, $E_{1,1} \in \mathscr{W}$. Then BAB, BA, BA^2 show that $E_{2,3}$, $E_{2,2}$, $E_{2,1} \in \mathscr{W}$. Finally, A^2B^2, A^2B, A^2 show that $E_{3,3}$, $E_{3,2}$, $E_{3,1} \in \mathscr{W}$. $\square$

Lemma 3. *Let* $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda \end{pmatrix}$ *where* $\lambda \neq 0$. *For any* 3×3 *matrix* A *such that* $\{A, B\}$ *is irreducible*, $M_3(\mathbb{C})$ *is spanned by words in* A *and* B *of length at most* 4.

Proof. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ and let $\mathscr{W}$ denote the span of the words in A and B of length at most 4. The nontrivial invariant subspaces of B are $\langle e_1 \rangle$, $\langle e_3 \rangle$, $\langle e_1, e_2 \rangle$, $\langle e_1, e_3 \rangle$ so, by irreducibility, at least one of each of the pairs $\{d, g\}$, $\{c, f\}$, $\{g, h\}$, $\{d, f\}$ is nonzero. Note that

$$B^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & a & c\lambda \\ 0 & d & f\lambda \\ 0 & g & i\lambda \end{pmatrix}, \quad BA = \begin{pmatrix} d & e & f \\ 0 & 0 & 0 \\ g\lambda & h\lambda & i\lambda \end{pmatrix},$$

$$B^2A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ g\lambda^2 & h\lambda^2 & i\lambda^2 \end{pmatrix}, \quad BAB = \begin{pmatrix} 0 & d & f\lambda \\ 0 & 0 & 0 \\ 0 & g\lambda & i\lambda^2 \end{pmatrix},$$

$$AB^2 = \begin{pmatrix} 0 & 0 & c\lambda^2 \\ 0 & 0 & f\lambda^2 \\ 0 & 0 & i\lambda^2 \end{pmatrix}, \quad B^2AB = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & g\lambda^2 & i\lambda^3 \end{pmatrix}.$$

Notice that the elementary matrices $E_{3,3}$, $E_{1,2}$ belong to $\mathscr{W}$ since $B^2 = \lambda^2 E_{3,3}$ and $\lambda B - B^2 = \lambda E_{1,2}$.

Case 1. d, f, g all nonzero. As we have noted, consideration of B and B^2 shows that $E_{3,3}, E_{1,2} \in \mathscr{W}$. A similar consideration of B^2AB then shows that $E_{3,2} \in \mathscr{W}$. Similarly, B^2A then shows that $E_{3,1} \in \mathscr{W}$. Then BAB shows that $E_{1,3} \in \mathscr{W}$, AB^2 that $E_{2,3} \in \mathscr{W}$, AB that $E_{2,2} \in \mathscr{W}$, BA that $E_{1,1} \in \mathscr{W}$, and finally A shows that $E_{2,1} \in \mathscr{W}$.

Case 2. $d = 0$. By irreducibility, $f \neq 0$ and $g \neq 0$. Note that

$$A^2 = \begin{pmatrix} * & * & * \\ fg & * & * \\ * & * & * \end{pmatrix}, \quad A^2B = \begin{pmatrix} 0 & * & * \\ 0 & fg & * \\ 0 & * & * \end{pmatrix}, \quad BA^2 = \begin{pmatrix} fg & * & * \\ 0 & 0 & 0 \\ * & * & * \end{pmatrix},$$

where a '$*$' denotes a scalar, possibly zero. As before, consideration of B and B^2 shows that $E_{3,3}, E_{1,2} \in \mathscr{W}$. Then B^2AB shows that $E_{3,2} \in \mathscr{W}$ and B^2A shows that $E_{3,1} \in \mathscr{W}$. Then BAB shows that $E_{1,3} \in \mathscr{W}$, AB^2 that $E_{2,3} \in \mathscr{W}$, A^2B that $E_{2,2} \in \mathscr{W}$, BA^2 that $E_{1,1} \in \mathscr{W}$, and finally A^2 shows that $E_{2,1} \in \mathscr{W}$.

Case 3. $f = 0$. By irreducibility, $c \neq 0$ and $d \neq 0$.

Sub-case 3(i). $f = 0$ and c, d, g are all nonzero. As before, B and B^2 show that $E_{3,3}, E_{1,2} \in \mathscr{W}$. Then AB^2 shows that $E_{1,3} \in \mathscr{W}$ and BAB shows that $E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{2,2} \in \mathscr{W}$, B^2A that $E_{3,1} \in \mathscr{W}$, BA that $E_{1,1} \in \mathscr{W}$, A that $E_{2,1} \in \mathscr{W}$, and finally A^2 shows that $E_{2,3} \in \mathscr{W}$.

Sub-case 3(ii). $f = g = 0$. By irreducibility, c, d, h are all nonzero. Note that

$$A^2 = \begin{pmatrix} * & * & * \\ * & * & * \\ dh & * & * \end{pmatrix}, \quad A^2B^2 = \begin{pmatrix} 0 & 0 & * \\ 0 & 0 & cd\lambda^2 \\ 0 & 0 & * \end{pmatrix},$$

where a '$*$' denotes a scalar, possibly zero. As usual, B and B^2 show that $E_{3,3}, E_{1,2} \in \mathscr{W}$. Then AB^2 shows that $E_{1,3} \in \mathscr{W}$ and B^2A shows that $E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{2,2} \in \mathscr{W}$, BA that $E_{1,1} \in \mathscr{W}$, A that $E_{2,1} \in \mathscr{W}$, A^2B^2 that $E_{2,3} \in \mathscr{W}$, and finally A^2 shows that $E_{3,1} \in \mathscr{W}$.

Case 4. $g = 0$. By irreducibility, d, h are both nonzero. We can suppose that $f \neq 0$, using Sub-case 3(ii). As usual, B and B^2 show that $E_{3,3}, E_{1,2} \in \mathscr{W}$. Then B^2A shows that $E_{3,2} \in \mathscr{W}$ and BAB shows that $E_{1,3} \in \mathscr{W}$. Then AB^2 shows that $E_{2,3} \in \mathscr{W}$, AB that $E_{2,2} \in \mathscr{W}$, BA that $E_{1,1} \in \mathscr{W}$, A that $E_{2,1} \in \mathscr{W}$, and finally A^2 shows that $E_{3,1} \in \mathscr{W}$.

This completes the proof. $\square$

Lemma 4. *Let* $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. *For any* 3×3 *matrix* A *such that* $\{A, B\}$ *is irreducible*, $M_3(\mathbb{C})$ *is spanned by words in* A *and* B *of length at most* 4.

Proof. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ and let $\mathscr{W}$ denote the span of the words in A and B of length at most 4. The nontrivial invariant subspaces of B are $\langle e_1 \rangle$, $\langle e_3 \rangle$, $\langle e_1, e_2 \rangle$, $\langle e_1, e_3 \rangle$ so, by irreducibility, at least one of each of the pairs $\{d, g\}$, $\{c, f\}$, $\{g, h\}$, $\{d, f\}$ is nonzero. Note that

$$B^2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} a & a+b & 0 \\ d & d+e & 0 \\ g & g+h & 0 \end{pmatrix},$$

$$BA = \begin{pmatrix} a+d & b+e & c+f \\ d & e & f \\ 0 & 0 & 0 \end{pmatrix},$$

$$B^2A = \begin{pmatrix} a+2d & b+2e & c+2f \\ d & e & f \\ 0 & 0 & 0 \end{pmatrix},$$

$$BAB = \begin{pmatrix} a+d & a+b+d+e & 0 \\ d & d+e & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad AB^2 = \begin{pmatrix} a & 2a+b & 0 \\ d & 2d+e & 0 \\ g & 2g+h & 0 \end{pmatrix},$$

$$BAB^2 = \begin{pmatrix} a+d & 2a+b+2d+e & 0 \\ d & 2d+e & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$B^2AB = \begin{pmatrix} a+2d & a+b+2d+2e & 0 \\ d & d+e & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notice that the elementary matrix $E_{1,2}$ belongs to $\mathscr{W}$ since $B^2 - B = E_{1,2}$.

Case 1. d, f, g all nonzero. As we have noted, $B^2 - B = E_{1,2} \in \mathscr{W}$. Since $BA(B^2 - B) = (a+d)E_{1,2} + dE_{2,2}$ it follows that $E_{2,2} \in \mathscr{W}$. The form of B now shows that $E_{1,1} \in \mathscr{W}$. Then BAB shows that $E_{2,1} \in \mathscr{W}$, $(B^2 - B)A$ that $E_{1,3} \in \mathscr{W}$, BA that $E_{2,3} \in \mathscr{W}$, $A(B^2 - B)$ that $E_{3,2} \in \mathscr{W}$, AB that $E_{3,1} \in \mathscr{W}$, and finally $A(B^2 - B)A$ shows that $E_{3,3} \in \mathscr{W}$.

Case 2. $d = 0$. By irreducibility, $f \neq 0$ and $g \neq 0$. As before, B and B^2 show that $E_{1,2} \in \mathscr{W}$. Then $(B^2 - B)A$ shows that $E_{1,3} \in \mathscr{W}$ and $A(B^2 - B)$ shows that $E_{3,2} \in \mathscr{W}$. Then $AB - BAB$ shows that $E_{3,1} \in \mathscr{W}$ and $B^2AB - B^2A$ that $E_{2,3} \in \mathscr{W}$.

If $i \neq 0$, $AB - A$ shows that $E_{3,3} \in \mathscr{W}$, $(B^2 - B)A^2$ that $E_{1,1} \in \mathscr{W}$, B that $E_{2,2} \in \mathscr{W}$ and finally $A(AB - BAB)$ shows that $E_{2,1} \in \mathscr{W}$.

If $i = 0$, $A(B^2 - B)A$ shows that $E_{3,3} \in \mathscr{W}$, $(B^2 - B)A^2$ that $E_{1,1} \in \mathscr{W}$, B that $E_{2,2} \in \mathscr{W}$, and finally, $A(AB - BAB)$ shows that $E_{2,1} \in \mathscr{W}$.

Case 3. $f = 0$. By irreducibility, $c \neq 0$ and $d \neq 0$.

Sub-case 3(i). $f = 0$ and c, d, g are all nonzero. As before, B and B^2 show that $E_{1,2} \in \mathscr{W}$. Then $(B^2 - B)A$ shows that $E_{1,1} \in \mathscr{W}$ and $BA(B^2 - B)$ shows that $E_{2,2} \in \mathscr{W}$. Then BAB shows that $E_{2,1} \in \mathscr{W}$, BA that $E_{1,3} \in \mathscr{W}$, $A(B^2 - B)$ that $E_{3,2} \in \mathscr{W}$, and AB^2 shows that $E_{3,1} \in \mathscr{W}$.

If $i \neq 0$, A then shows that $E_{3,3} \in \mathscr{W}$ and finally AB^2A shows that $E_{2,3} \in \mathscr{W}$.

If $i = 0$, BA^2 shows that $E_{3,2} \in \mathscr{W}$ and finally A^2 that $E_{3,3} \in \mathscr{W}$.

Sub-case 3(ii). $f = g = 0$. By irreducibility, c, d, h are all nonzero. As in the preceding sub-case, B, B^2 and $(B^2 - B)A$, $BA(B^2 - B)$, BAB, BA show successively that $E_{1,2}$ and $E_{1,1}$, $E_{2,2}$, $E_{2,1}$, $E_{1,3} \in \mathscr{W}$, respectively. Then $B^2A - AB^2$ shows that $E_{3,2} \in \mathscr{W}$.

If $i \neq 0$, A then shows that $E_{3,3} \in \mathscr{W}$ and A^2B that $E_{3,1} \in \mathscr{W}$. Finally A^2 shows that $E_{2,3} \in \mathscr{W}$.

If $i = 0$, A^2B then shows that $E_{3,1} \in \mathscr{W}$ and A^2 that $E_{2,3} \in \mathscr{W}$. Finally A^3 shows that $E_{3,3} \in \mathscr{W}$.

Case 4. $g = 0$. By irreducibility, d, h are both nonzero. We can suppose that $f \neq 0$, using Sub-case 3(ii). As usual, B and B^2 show that $E_{1,2} \in \mathscr{W}$. Then $A(B^2 - B)$ shows that $E_{2,2} \in \mathscr{W}$ and B shows that $E_{1,1} \in \mathscr{W}$. Then BAB shows that $E_{2,1} \in \mathscr{W}$, $(B^2 - B)A$ that $E_{1,3} \in \mathscr{W}$, BA that $E_{2,3} \in \mathscr{W}$, and AB^2 that $E_{3,2} \in \mathscr{W}$.

If $i \neq 0$, $ABAB$ then shows that $E_{3,1} \in \mathscr{W}$ and finally A shows that $E_{3,3} \in \mathscr{W}$.

If $i = 0$, A^2B then shows that $E_{3,1} \in \mathscr{W}$ and finally A^2 shows that $E_{3,3} \in \mathscr{W}$.

This completes the proof. □

Lemma 5. *Let* $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}$ *where* $\lambda \neq 0, 1$. *For any* 3×3 *matrix* A *such that* $\{A, B\}$ *is irreducible,* $M_3(\mathbb{C})$ *is spanned by words in* A *and* B *of length at most* 4.

Proof. Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ and let $\mathscr{W}$ denote the span of the words in A and B of length at most 4. The nontrivial invariant subspaces of B are $\langle e_1 \rangle$, $\langle e_2 \rangle$, $\langle e_3 \rangle$, $\langle e_1, e_2 \rangle$, $\langle e_1, e_3 \rangle$, $\langle e_2, e_3 \rangle$ so, by irreducibility, at least one of each of the pairs $\{d, g\}$, $\{b, h\}$, $\{c, f\}$, $\{g, h\}$, $\{d, f\}$, $\{b, c\}$ is nonzero. This leads to 18 cases to be considered. However, only 11 cases need to be considered in detail because of adjoints. For example, the proof for the case where $c = d = h = 0$ and b, f, g are all nonzero will follow from that for the case where $b = f = g = 0$ and c, d, h are all nonzero. More precisely, if the words in A^* and B^* of length at most 4 span $M_3(\mathbb{C})$, then the words in A and B of length at most 4 span $M_3(\mathbb{C})$. Note that

$$B^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & b & c\lambda \\ 0 & e & f\lambda \\ 0 & h & i\lambda \end{pmatrix},$$

$$BA = \begin{pmatrix} 0 & 0 & 0 \\ d & e & f \\ g\lambda & h\lambda & i\lambda \end{pmatrix}, \quad B^2A = \begin{pmatrix} 0 & 0 & 0 \\ d & e & f \\ g\lambda^2 & h\lambda^2 & i\lambda^2 \end{pmatrix},$$

$$BAB = \begin{pmatrix} 0 & 0 & 0 \\ 0 & e & f\lambda \\ 0 & h\lambda & i\lambda^2 \end{pmatrix}, \quad AB^2 = \begin{pmatrix} 0 & b & c\lambda^2 \\ 0 & e & f\lambda^2 \\ 0 & h & i\lambda^2 \end{pmatrix},$$

$$B^2AB = \begin{pmatrix} 0 & 0 & 0 \\ 0 & e & f\lambda \\ 0 & h\lambda^2 & i\lambda^3 \end{pmatrix}, \quad BAB^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & e & f\lambda^2 \\ 0 & h\lambda & i\lambda^3 \end{pmatrix}.$$

Notice that the elementary matrices $E_{2,2}$, $E_{3,3}$ belong to $\mathscr{W}$ since $B^2 - B = (\lambda^2 - \lambda)E_{3,3}$ and $B^2 - \lambda B = (1-\lambda)E_{2,2}$.

Case 1. b, c, d, f, g, h all nonzero. As we have noted, consideration of B and B^2 shows that $E_{2,2}, E_{3,3} \in \mathscr{W}$. A similar consideration of BAB and BAB^2 then shows that $E_{2,3}, E_{3,2} \in \mathscr{W}$. Similarly, AB and AB^2 then show that $E_{1,2}, E_{1,3} \in \mathscr{W}$. Then BA and B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. Finally, if $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$. If $a = 0$, since then

$$A^2 = \begin{pmatrix} bd + cg & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \quad \text{and} \quad ABA = \begin{pmatrix} bd + cg\lambda & * & * \\ * & * & * \\ * & * & * \end{pmatrix},$$

where '$*$' denotes a scalar, possibly zero, and where not both $bd + cg$ and $bd + cg\lambda$ can be zero, $E_{1,1} \in \mathscr{W}$.

Case 2. $b = 0$ and c, d, f, g, h all nonzero. As before, B and B^2 show that $E_{2,2}$, $E_{3,3} \in \mathscr{W}$. Then BAB and BAB^2 show that $E_{2,3}, E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{1,3} \in \mathscr{W}$. Then BA and B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. Finally, if $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$ and A^2 then shows that $E_{1,2} \in \mathscr{W}$. If $a = 0$, A^2B shows that $E_{1,2} \in \mathscr{W}$ then A^2 shows that $E_{1,1} \in \mathscr{W}$.

Case 3. $c = 0$ and b, d, f, g, h all nonzero. Again, B and B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$ and BAB and BAB^2 then show that $E_{2,3}, E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{1,2} \in \mathscr{W}$. Then BA and B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. Finally, if $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$ and A^2 then shows that $E_{1,3} \in \mathscr{W}$. If $a = 0$, A^2B shows that $E_{1,3} \in \mathscr{W}$ then A^2 shows that $E_{1,1} \in \mathscr{W}$.

Case 4. $f = 0$ and b, c, d, g, h all nonzero. Again, B and B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB shows that $E_{3,2} \in \mathscr{W}$. Then AB and AB^2 show that $E_{1,2}, E_{1,3} \in \mathscr{W}$. If $a \neq 0$, A, BA, B^2A show that $E_{1,1}, E_{2,1}, E_{3,1} \in \mathscr{W}$ and A^2 then shows that $E_{2,3} \in \mathscr{W}$. If $a = 0$, A, BA show that $E_{2,1}, E_{3,1} \in \mathscr{W}$ then A^2B shows that $E_{2,3} \in \mathscr{W}$ and A^2 or ABA shows that $E_{1,1} \in \mathscr{W}$.

Case 5. $b = d = 0$ and c, f, g, h all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB, B^2AB show that $E_{2,3}, E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{1,3} \in \mathscr{W}$

and BA shows that $E_{3,1} \in \mathscr{W}$ and A^2B shows that $E_{1,2} \in \mathscr{W}$. If $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$, and A^2 then shows that $E_{2,1} \in \mathscr{W}$. If $a = 0$, BA^2 shows that $E_{2,1} \in \mathscr{W}$ then A^2 shows that $E_{1,1} \in \mathscr{W}$.

Case 6. $b = f = 0$ and c, d, g, h all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB shows that $E_{3,2} \in \mathscr{W}$ and BA^2B shows that $E_{2,3} \in \mathscr{W}$. Then AB shows that $E_{1,3} \in \mathscr{W}$ and BA, B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. If $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$, and A^2 then shows that $E_{1,2} \in \mathscr{W}$. If $a = 0$, A^2B shows that $E_{1,2} \in \mathscr{W}$ then A^2 shows that $E_{1,1} \in \mathscr{W}$.

Case 7. $b = g = 0$ and c, d, f, h all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB, BAB^2 show that $E_{2,3}, E_{3,2} \in \mathscr{W}$. Then AB shows that $E_{1,3} \in \mathscr{W}$, A^2B shows that $E_{1,2} \in \mathscr{W}$ and BA shows that $E_{2,1} \in \mathscr{W}$. If $a \neq 0$, A then shows that $E_{1,1} \in \mathscr{W}$, and A^2 that $E_{3,1} \in \mathscr{W}$. If $a = 0$, A^2 shows that $E_{3,1} \in \mathscr{W}$ and A^2BA shows that $E_{1,1} \in \mathscr{W}$.

Case 8. $c = g = 0$ and b, d, f, h all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB, BAB^2 show that $E_{2,3}, E_{3,2} \in \mathscr{W}$ and AB then shows that $E_{1,2} \in \mathscr{W}$, A^2B that $E_{1,3} \in \mathscr{W}$ and BA that $E_{2,1} \in \mathscr{W}$. If $a \neq 0$, A then shows that $E_{1,1} \in \mathscr{W}$, and A^2 that $E_{3,1} \in \mathscr{W}$. If $a = 0$, BA^2 shows that $E_{3,1} \in \mathscr{W}$ and A^2 shows that $E_{1,1} \in \mathscr{W}$.

Case 9. $c = h = 0$ and b, d, f, g all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB shows that $E_{2,3} \in \mathscr{W}$, BA^2B that $E_{3,2} \in \mathscr{W}$ and AB^2 that $E_{1,2} \in \mathscr{W}$. Then BA, B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. If $a \neq 0$, A then shows that $E_{1,1} \in \mathscr{W}$, and A^2 that $E_{1,3} \in \mathscr{W}$. If $a = 0$, A^2B shows that $E_{1,3} \in \mathscr{W}$ and A^2 then shows that $E_{1,1} \in \mathscr{W}$.

Case 10. $f = h = 0$ and b, c, d, g all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then AB, AB^2 show that $E_{1,2}, E_{1,3} \in \mathscr{W}$. Then BA, B^2A show that $E_{2,1}, E_{3,1} \in \mathscr{W}$. If $a \neq 0$, A shows that $E_{1,1} \in \mathscr{W}$ and then A^2, BA^2 show that $E_{2,3}, E_{3,2} \in \mathscr{W}$. If $a = 0$, A^2B, A^2B^2 show that $E_{2,3}, E_{3,2} \in \mathscr{W}$, then ABA or AB^2A shows that $E_{1,1} \in \mathscr{W}$.

Case 11. $b = f = g = 0$ and c, d, h all nonzero. As usual, B, B^2 show that $E_{2,2}, E_{3,3} \in \mathscr{W}$. Then BAB shows that $E_{3,2} \in \mathscr{W}$, AB that $E_{1,3} \in \mathscr{W}$, BA that $E_{2,1} \in \mathscr{W}$, BA^2B that $E_{2,3} \in \mathscr{W}$, A^2B that $E_{1,2} \in \mathscr{W}$, BA^2 that $E_{3,1} \in \mathscr{W}$ and, finally, A^2BA that $E_{1,1} \in \mathscr{W}$.

This completes the proof. □

The following is the main result of this section.

Theorem 1. *Let* $\dim \mathscr{H} = 3$ *and let* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible. Then* $\mathscr{B}(\mathscr{H})$ *is spanned by words in* A *and* B *of length at most* 4. *Consequently* $\mathrm{msl}(A, B) = 3$ *or* 4.

Proof. We can suppose that $\mathscr{H} = \mathbb{C}^3$ and that $A, B \in M_3(\mathbb{C})$. Since neither A nor B can be a multiple of the identity, the degrees of their minimum polynomials must be at least 2. These degrees cannot both be equal to 2 since every pair of quadratic transformations on a space of (finite) dimension greater than 2 has a common non-trivial invariant subspace by [2, Theorem 2]. We may assume that the degree of the minimum polynomial of B is 3.

Case 1. B has only one eigenvalue, α say. Then $(B - \alpha I)^3 = 0$ and B is similar to $\begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix} = \alpha I + J$, where J is the strictly upper triangular elementary Jordan matrix. We may suppose that $B = \alpha I + J$. By Lemma 2, $M_3(\mathbb{C})$ is spanned by words in A and J of length at most 4. If $\alpha \neq 0$, then B is invertible so the identity can be written as a cubic polynomial in B with no constant term (obtained by expanding $(B - \alpha I)^3$). Hence the span of the words in A and B of length at most 4 contains the identity. Therefore, it also contains every word in A and J of length at most 4. (If J is replaced by $B - \alpha I$ in the word, and the factors in $(B - \alpha I)^2$, $(B - \alpha I)^3$, $(B - \alpha I)^4$ are expanded, the result is a linear combination of words of length at most 4 in A and B plus a multiple of the identity.) Consequently, $M_3(\mathbb{C})$ is spanned by words in A and B of length at most 4.

Case 2. B has precisely two distinct eigenvalues, α and β, say. Then we may suppose that $(B - \alpha I)^2(B - \beta I) = 0$, and that $B = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}$. If $\alpha = 0$, the desired result follows from Lemma 3. If $\alpha \neq 0$ and $\beta \neq 0$, then B is invertible. Consequently, the identity can be written as a cubic polynomial in B with no constant term (this polynomial is obtained by expanding $(B - \alpha I)^2(B - \beta I)$). By Lemma 3 the words in A and $B - \alpha I$ of length at most 4 span $M_3(\mathbb{C})$ and every such word belongs to the span of the words in A, B of length at most 4 and I. Since, as noted, I belongs to the span of the words in A, B of length at most 4, it follows that the latter equals $M_3(\mathbb{C})$.

This leaves the case where $\alpha \neq 0$ and $\beta = 0$. Then $\frac{B}{\alpha}$ is similar to $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and we may suppose that B equals the latter matrix. The desired result then follows from Lemma 4.

Case 3. B has three distinct eigenvalues. In this case we may suppose that $B = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$ where α, β and γ are distinct scalars.

If B is not invertible, we may suppose that $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$. Then $\frac{B}{\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}$ where $\lambda = \gamma/\beta \neq 0, 1$. The desired result now follows from Lemma 5.

If B is invertible then $\frac{B-\alpha I}{\beta-\alpha} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}$ where $\lambda = (\gamma - \alpha)/(\beta - \alpha) \neq 0, 1$. By Lemma 5, the words in A and $B - \alpha I$ of length at most 4 span $M_3(\mathbb{C})$. Since the identity matrix can be written as a cubic polynomial in B with no constant term, and since every word in A and $B - \alpha I$ is a span of words in A, B of no greater length and I, the desired result follows. □

The following example shows that $\mathrm{msl}(A, B) = 3$ is possible on three-dimensional space.

Example 1. Put $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. The nontrivial invariant subspaces of B are $\langle e_1 \rangle$ and $\langle e_1, e_2 \rangle$ so clearly $\{A, B\}$ is irreducible. It is easily checked that

$$\{A, B, A^2, AB, BA, B^2, A^2B, AB^2, BA^2\}$$

is a basis for $M_3(\mathbb{C})$. (In fact $E_{1,3} = B^2$, $E_{2,3} = AB^2 - B^2$, $E_{3,2} = A - BA^2$, $E_{1,2} = B - AB^2 + B^2$, $E_{3,1} = A^2 - BA^2$, $E_{2,2} = A^2B - B + AB^2 - B^2 - A + BA^2$, $E_{3,3} = AB - A^2B + A - BA^2$, $E_{1,1} = BA - A^2B + B - AB^2 + B^2 + A - BA^2$, $E_{2,1} = A^2B - BA - B + AB^2 - B^2 - A + 2BA^2$.) Consequently $\mathrm{msl}(A, B) = 3$.

An example showing that $\mathrm{msl}(A, B) = 4$ can occur on three-dimensional space can be obtained by taking B as in Example 1 and taking $A = (B^*)^2$. See Example 2 below for the details, putting $n = 3$.

3. Unicellular case

The main result of this section (Theorem 2) shows that $\mathrm{msl}(A, B) \leqslant 2n - 2$ whenever B is a unicellular transformation on $\mathscr{H}$ and $\{A, B\}$ is irreducible, where $n = \dim \mathscr{H} \geqslant 2$. Recall that a linear transformation is unicellular if its set of invariant subspaces is a chain, that is, is totally ordered by inclusion. Some remarks on the proof of the theorem that follows may make it more easily understandable.

Remarks. (1) In the proof we may suppose that $\mathscr{H} = \mathbb{C}^n$. Since B is unicellular it is similar to a scalar translate of the strictly upper triangular elementary Jordan matrix J. We can suppose that $B = \alpha I + J$, for some scalar α. In fact, it is enough to consider the case where $\alpha = 0$. For if $\alpha \neq 0$, the equation $(B - \alpha I)^n = 0$ shows that I can be expressed as a polynomial in B of degree n, with no constant term. Thus the span of the words in A and B of length at most $2n - 2$ contains I. It follows that it also contains every word in A and J of length at most $2n - 2$ since every word in A and $J = B - \alpha I$ is a linear combination of words in A and B of no greater length plus a scalar multiple of the identity. Thus if $\mathrm{msl}(A, J) \leqslant 2n - 2$ then $\mathrm{msl}(A, B) \leqslant 2n - 2$.

(2) Consider the operation of pre-multiplying by B. If C is any $n \times n$ matrix, BC is obtained from C by deleting the first row of C and inserting a row of zeros as the last row (in other words, pre-multiplication by B lifts C up one row and introduces a row of zeros at the bottom). Post-multiplication by B has a similar description. The matrix CB is obtained from C by deleting the last column of C and inserting a column of zeros as the first column (post-multiplication by B moves C one column to the right and introduces a column of zeros on the left). Thus, for integers r, s where $0 \leqslant r, s \leqslant n - 1$, $B^r C B^s$ is obtained by lifting C up r rows, introducing r rows of zeros at the bottom, then moving the resulting matrix s columns to the right, introducing s columns of zeros on the left. Schematically

$$B^r C B^s = \begin{pmatrix} 0_{n-r,s} & C_{n-r,n-s} \\ 0_{r,s} & 0_{r,n-s} \end{pmatrix},$$

where $0_{u,v}$ denotes the $u \times v$ zero matrix and where $C_{u,v}$ denotes the bottom left hand corner sub-matrix of C of size $u \times v$.

(3) In the proof we exhibit a set of n^2 words in A and B and show that it is a basis for $M_n(\mathbb{C})$ by showing that its span contains every elementary matrix $E_{i,j}$. We repeatedly use the fact that, if the span of a set $\mathscr{B}$ of matrices contains each elementary matrix in the set $\{E_{i,j} : (i, j) \in \mathscr{S}\}$ where $\mathscr{S}$ is some subset of $\{1, 2 \dots, n\} \times \{1, 2 \dots, n\}$, and it contains a matrix the positions of whose nonzero entries form a subset of $\mathscr{S}$ with one exception, say (u, v), then $E_{u,v}$ also belongs to the span of $\mathscr{B}$.

Theorem 2. *Let* $\dim \mathscr{H} = n \geqslant 2$ *and let* $B \in \mathscr{B}(\mathscr{H})$ *be unicellular. For any transformation* $A \in \mathscr{B}(\mathscr{H})$ *such that* $\{A, B\}$ *irreducible,* $\mathscr{B}(\mathscr{H})$ *is spanned by words in* A *and* B *of length at most* $2n - 2$. *Consequently* $\mathrm{msl}(A, B) \leqslant 2n - 2$.

Proof. We may suppose that $\mathscr{H} = \mathbb{C}^n$ and that B is the $n \times n$ strictly upper triangular elementary Jordan matrix (see above remarks). Let $A = (a_{i,j}) \in M_n(\mathbb{C})$ with $\{A, B\}$ irreducible.

For each $1 \leqslant k \leqslant n$ let d_k be the largest integer satisfying $a_{d_k,k} \neq 0$, taking $d_k = 0$ if the kth column of A is zero (so d_k measures the 'depth' of the kth column of A). Since the nontrivial invariant subspaces of B are $\langle e_1, e_2, \dots, e_j\rangle$, $1 \leqslant j \leqslant n - 1$, for each $1 \leqslant j \leqslant n - 1$ at least one d_k, $1 \leqslant k \leqslant j$ satisfies $d_k \geqslant j + 1$. For $1 \leqslant j \leqslant n - 1$ let m_j be the smallest integer satisfying $d_{m_j} \geqslant j + 1$. Then $d_k \leqslant j$ if $k < m_j$. Clearly $1 = m_1 \leqslant m_2 \leqslant \cdots \leqslant m_{n-1}$.

Put $p_0 = n$, $p_1 = m_{p_0-1}$, $p_2 = m_{p_1-1}$ and, in general, put $p_{i+1} = m_{p_i-1}$ whenever $p_i \geqslant 2$. Then $p_{i+1} < p_i$ since $m_j \leqslant j$ when $j = p_i - 1$. Let t be the integer satisfying $p_t = 1$. Note that $p_i \leqslant d_{p_{i+1}} < p_{i-1}$ for $1 \leqslant i \leqslant t - 1$. The former inequality follows from the fact that $d_{m_j} \geqslant j + 1$ (take $j = p_i - 1$) and the latter from the fact that $d_k \leqslant j$ if $k < m_j$ (take $j = p_{i-1} - 1$). We show how the natural numbers $\{d_{p_{i+1}} - p_i : 1 \leqslant i \leqslant t-1\}$ determine a basis for $M_n(\mathbb{C})$ consisting of words in A and B of length no more than $2n-2$.

We have

$$\begin{aligned} 1 = p_t &< p_{t-1} \leqslant d_{p_t} < p_{t-2} \leqslant d_{p_{t-1}} < p_{t-3} \leqslant \cdots \\ &< p_2 \leqslant d_{p_3} < p_1 \leqslant d_{p_2} < p_0 = n. \end{aligned}$$

Define the words $W_1, W_2, \ldots, W_t$ in A and B inductively by

$$W_1 = A \quad \text{and} \quad W_{i+1} = W_i B^{d_{p_{i+1}} - p_i} A, \quad 1 \leqslant i \leqslant t-1.$$

Then $\text{length}(W_1) = 1$ and, for $1 \leqslant i \leqslant t-1$,

$$\text{length}(W_{i+1}) = \text{length}(W_i) + d_{p_{i+1}} - p_i + 1 = \sum_{k=1}^{i} (d_{p_{k+1}} - p_k) + i + 1.$$

Claim. $\text{length}(W_{i+1}) \leqslant d_{p_2} - p_i + 2$ for $1 \leqslant i \leqslant t-1$.

It is obvious that $\text{length}(W_2) \leqslant d_{p_2} - p_1 + 2$. For $2 \leqslant i \leqslant t-1$ we have

$$\begin{aligned} &\text{length}(W_{i+1}) + \sum_{k=1}^{i-1} (p_k - d_{p_{k+2}}) \\ &\quad = \sum_{k=1}^{i-1} (d_{p_{k+1}} - p_k + p_k - d_{p_{k+2}}) + d_{p_{i+1}} - p_i + i + 1 \\ &\quad = d_{p_2} - d_{p_{i+1}} + d_{p_{i+1}} - p_i + i + 1 = d_{p_2} - p_i + i + 1. \end{aligned}$$

But $\sum_{k=1}^{i-1} (p_k - d_{p_{k+2}}) \geqslant i-1$ (since $p_k > d_{p_{k+2}}$ for $1 \leqslant k \leqslant i-1$), so the claim follows.

Note that the first nonzero entry in the last row of W_1 (counting from left to right) occurs in column p_1. Also, the first nonzero entry in the last row of W_2 occurs in column p_2. In general, the first nonzero entry in the last row of W_i occurs in column p_i.

Claim 0. The following set $\mathscr{B}$ of matrices is a basis for $M_n(\mathbb{C})$

$$\begin{aligned} \mathscr{B} = \{B^{n-1}\} &\cup \big\{ B^r W_1 B^s : 0 \leqslant r \leqslant n-1, 0 \leqslant s \leqslant n - p_1, \\ &\qquad \text{and } r + s < 2n - 1 - p_1 \big\} \\ &\cup \bigcup_{i=1}^{t-1} \big\{ B^r W_{i+1} B^s : 0 \leqslant r \leqslant n-1, 0 \leqslant s \leqslant p_i - p_{i+1} - 1 \big\}. \end{aligned}$$

Note that each word comprising $\mathscr{B}$ has length no more than $2n-2$. Indeed, $\text{length}(B^r W_1 B^s) = r + s + 1$ and $r + s + 1 \leqslant 2n - 1 - p_1 \leqslant 2n - 2$ if $r + s < 2n - 1 - p_1$. Also, for $1 \leqslant i \leqslant t-1$ and $0 \leqslant r \leqslant n-1$, $0 \leqslant s \leqslant p_i - p_{i+1} - 1$,

$$\begin{aligned}\text{length}(B^r W_{i+1} B^s) &= \text{length}(W_{i+1}) + r + s \\ &\leqslant \text{length}(W_{i+1}) + n - 2 + p_i - p_{i+1} \\ &\leqslant d_{p_2} - p_i + 2 + n - 2 + p_i - p_{i+1} \\ &= n + d_{p_2} - p_{i+1} \\ &\leqslant 2n - 2,\end{aligned}$$

since $d_{p_2} \leqslant n - 1$ and $p_{i+1} \geqslant 1$.

Note also that $\mathscr{B}$ is comprised of n^2 matrices since

$$\begin{aligned}&1 + n(n - p_1 + 1) - 1 + \sum_{i=1}^{t-1} n(p_i - p_{i+1}) \\ &\quad = n\left(n - p_1 + 1 + \sum_{i=1}^{t-1}(p_i - p_{i+1})\right) \\ &\quad = n(n + 1 - p_t) = n^2.\end{aligned}$$

Claim 0 is proved by showing that each of the elementary matrices $E_{r,s}$, $1 \leqslant r, s \leqslant n$, belongs to the span $\langle \mathscr{B} \rangle$ of $\mathscr{B}$. Clearly $E_{1,n} = B^{n-1} \in \langle \mathscr{B} \rangle$.

Claim 1. $\{E_{1,s} : 1 \leqslant s \leqslant n\} \subseteq \langle \mathscr{B} \rangle$.

We prove this claim in t steps.

Step 0. We show that $\{E_{1,s} : p_1 \leqslant s \leqslant n - 1\} \subseteq \langle \mathscr{B} \rangle$. For $0 \leqslant s \leqslant n - 1 - p_1$, the matrix $B^{n-1} W_1 B^s$ belongs to $\mathscr{B}$ and is equal to a nonzero scalar multiple of $E_{1,s+p_1}$ plus a linear combination of $\{E_{1,q} : s + p_1 + 1 \leqslant q \leqslant n\}$. Thus if $\{E_{1,q} : s + p_1 + 1 \leqslant q \leqslant n\} \subseteq \langle \mathscr{B} \rangle$ then $E_{1,s+p_1} \in \langle \mathscr{B} \rangle$. Taking $s = n - 1 - p_1$, and using the fact that $E_{1,n} \in \langle \mathscr{B} \rangle$, gives that $E_{1,n-1} \in \langle \mathscr{B} \rangle$. The desired result now follows by induction.

Step i ($1 \leqslant i \leqslant t - 1$). We show that $\{E_{1,s} : p_{i+1} \leqslant s \leqslant p_i - 1\} \subseteq \langle \mathscr{B} \rangle$. For $0 \leqslant s \leqslant p_i - p_{i+1} - 1$, the matrix $B^{n-1} W_{i+1} B^s$ belongs to $\mathscr{B}$ and is equal to a nonzero scalar multiple of $E_{1,s+p_{i+1}}$ plus a linear combination of $\{E_{1,q} : s + p_{i+1} + 1 \leqslant q \leqslant n\}$. Taking $s = p_i - p_{i+1} - 1$, and using the results of Steps 0 to $i - 1$, gives that $E_{1,p_i-1} \in \langle \mathscr{B} \rangle$. The desired result now follows by induction.

These t steps together prove Claim 1.

Claim 2. $\{E_{2,s} : 1 \leqslant s \leqslant n\} \subseteq \langle \mathscr{B} \rangle$.

We prove this claim in t steps.

Step 0. We show that $\{E_{2,s} : p_1 \leqslant s \leqslant n\} \subseteq \langle \mathscr{B} \rangle$. For $0 \leqslant s \leqslant n - p_1$, the matrix $B^{n-2} W_1 B^s$ belongs to $\mathscr{B}$ and is equal to a nonzero scalar multiple of $E_{2,s+p_1}$ plus a linear combination of $\{E_{2,q} : s + p_1 + 1 \leqslant q \leqslant n\} \cup \{E_{1,q} : 1 \leqslant q \leqslant n\}$ (take the first set to be empty if $s = n - p_1$). Taking $s = n - p_1$, and using the result of Claim 1, gives that $E_{2,n} \in \langle \mathscr{B} \rangle$. The desired result now follows by induction.

Step i ($1 \leqslant i \leqslant t-1$). We show that $\{E_{2,s} : p_{i+1} \leqslant s \leqslant p_i - 1\} \subseteq \langle\mathscr{B}\rangle$. For $0 \leqslant s \leqslant p_i - p_{i+1} - 1$, the matrix $B^{n-2}W_{i+1}B^s$ belongs to $\mathscr{B}$ and is equal to a non-zero scalar multiple of $E_{2,s+p_{i+1}}$ plus a linear combination of $\{E_{2,q} : s + p_{i+1} + 1 \leqslant q \leqslant n\} \cup \{E_{1,q} : 1 \leqslant q \leqslant n\}$. Taking $s = p_i - p_{i+1} - 1$, and using the results of Steps 0 to $i-1$ and Claim 1, gives that $E_{2,p_i-1} \in \langle\mathscr{B}\rangle$. The desired result now follows by induction.

These t steps together prove Claim 2.

In a similar way we can prove successively, and each in t steps, the claims Claim r ($3 \leqslant r \leqslant n$): $\{E_{r,s} : 1 \leqslant s \leqslant n\} \subseteq \langle\mathscr{B}\rangle$.

Together Claims 1 to n prove Claim 0 and the sproof is complete. □

In the statement of the preceding theorem the upper bound of $2n-2$ is sharp, in the sense that, on every space $\mathscr{H}$ of dimension $n \geqslant 2$ there exists a unicellular transformation B and a transformation A with $\{A, B\}$ irreducible such that $\mathrm{msl}(A, B) = 2n-2$, that is, such that the words in A and B of length at most $2n-3$ do not span $\mathscr{B}(\mathscr{H})$. For $n = 2$ this is obvious. For $n \geqslant 3$ we have the following.

Example 2. Let $n \geqslant 3$, let B be the $n \times n$ strictly upper triangular elementary Jordan matrix and let $A = (B^*)^{n-1}$. Then the n, 1-entry of A is one and all other entries are zero. Every word in A and B maps each basis vector e_j to a basis vector or to the zero vector. Let W be a word in A and B of length at most $2n-3$. We show that $(We_n|e_2) = (We_{n-1}|e_1)$. This is certainly true if both $(We_n|e_2)$ and $(We_{n-1}|e_1)$ are zero.

Suppose that $(We_n|e_2) \neq 0$. Then $We_n = e_2$. If W contained a factor of A it would be of the form $W = VAB^{n-1}$ where V is a word in A and B satisfying $Ve_n = e_2$. But then V would have length at least $n-2$ and W would have length at least $2n-2$. This contradiction shows that W contains no factors of A, so $W = B^{n-2}$. Thus $(We_n|e_2) = (We_{n-1}|e_1) = 1$.

Suppose that $(We_{n-1}|e_1) \neq 0$. Then $We_{n-1} = e_1$. If W contained a factor of A it would be of the form $W = VAB^{n-2}$ where V is a word in A and B satisfying $Ve_n = e_1$. But then V would have length at least $n-1$ and W would have length at least $2n-2$. This contradiction shows that, once again, W contains no factors of A, so $W = B^{n-2}$. Thus $(We_n|e_2) = (We_{n-1}|e_1) = 1$.

Since $(We_n|e_2) = (We_{n-1}|e_1)$ for every word W in A and B of length at most $2n-3$, the same equality holds for every element in the span of such words. Hence the span of such words cannot be $M_n(\mathbb{C})$. By Theorem 2 it follows that $\mathrm{msl}(A, B) = 2n-2$.

We next show that, on every space $\mathscr{H}$ of dimension $n \geqslant 2$ there exists a unicellular transformation B and a transformation A with $\{A, B\}$ irreducible such that $\mathrm{msl}(A, B) = n$. In this we can suppose, as usual, that $\mathscr{H} = \mathbb{C}^n$ and that B is the strictly upper triangular elementary Jordan matrix. Example 1 above provides an example for the case $n = 3$. For $n \neq 3$ the following two propositions provide an example.

Proposition 3. *Let $n \geqslant 2, n \neq 3$, let B be the $n \times n$ strictly upper triangular elementary Jordan matrix, and let $A = B^*$. Then $M_n(\mathbb{C})$ is spanned by words in A and B of length at most n. Consequently* $\operatorname{msl}(A, B) \leqslant n$.

Proof. It is clear that $\{A, B\}$ is irreducible. By Theorem 2 (or Proposition 2) the result is true for $n = 2$. Let $n \geqslant 4$. For $0 \leqslant p, q \leqslant n-1$ and $1 \leqslant r, s \leqslant n$ we have

$$B^p e_r = \begin{cases} e_{r-p} & \text{if } r-p \geqslant 1, \\ 0 & \text{if } r-p \leqslant 0 \end{cases} \quad \text{and} \quad A^q e_s = \begin{cases} e_{s+q} & \text{if } s+q \leqslant n, \\ 0 & \text{if } s+q \geqslant n+1. \end{cases}$$

For any $n \times n$ matrix $C = (c_{i,j})$ and any integer $t \in \{0, \pm 1, \pm 2, \ldots, \pm(n-1)\}$ define the t-diagonal of C to be the set of entries $\{c_{i,j} : i - j = t\}$. For each integer $-(n-1) \leqslant t \leqslant n-1$ it is easily verified that each of the matrices belonging to

$$\left\{A^{p+t} B^p : 0 \leqslant p \leqslant n-1, 0 \leqslant p+t \leqslant n-1\right\}$$
$$\cup \left\{B^p A^{p+t} : 0 \leqslant p \leqslant n-1, 0 \leqslant p+t \leqslant n-1\right\}$$

has zero entries off its t-diagonal. We show that, for each integer $0 < |t| \leqslant n-1$, this set of matrices contains an appropriate basis for the subspace of matrices having zeros off the t-diagonal. The latter has dimension $n - |t|$, of course. The case $t = 0$ is treated separately.

Case 1. $t = 0$.

Sub-case 1(i). n even, say $n = 2m$ (with $m \geqslant 2$). We show that

$$\{AB, A^2B^2, \ldots, A^mB^m\} \cup \{BA, B^2A^2, \ldots, B^mA^m\}$$

is linearly independent (and so forms a basis for the subspace of diagonal matrices). Suppose that

$$\sum_{p=1}^{m} \alpha_p A^p B^p = \sum_{q=1}^{m} \beta_q B^q A^q.$$

Applying both sides of this equation to e_r, $2 \leqslant r \leqslant m$, we get

$$\sum_{p=1}^{r-1} \alpha_p = \sum_{q=1}^{m} \beta_q \quad \text{for } 2 \leqslant r \leqslant m.$$

Applying them to e_r, $m+1 \leqslant r \leqslant 2m-1$, gives

$$\sum_{p=1}^{m} \alpha_p = \sum_{q=1}^{2m-r} \beta_q, \quad \text{for } m+1 \leqslant r \leqslant 2m-1.$$

Finally, applying them to e_1 and e_{2m} gives

$$\sum_{p=1}^{m} \alpha_p = \sum_{q=1}^{m} \beta_q = 0.$$

It follows that $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \beta_1 = \beta_2 = \cdots = \beta_m = 0$.

Sub-case 1(ii). n odd, say $n = 2m + 1$ (with $m \geqslant 2$). The matrix BA^2B is diagonal. We show that

$$\{AB, A^2B^2, \ldots, A^mB^m\} \cup \{BA, B^2A^2, \ldots, B^mA^m\} \cup \{BA^2B\}$$

is linearly independent. Suppose that

$$\sum_{p=1}^{m} \alpha_p A^p B^p = \sum_{q=1}^{m} \beta_q B^q A^q + \gamma BA^2B.$$

Applying both sides to e_r, $2 \leqslant r \leqslant m$, we get

$$\sum_{p=1}^{r-1} \alpha_p = \sum_{q=1}^{m} \beta_q + \gamma \quad \text{for } 2 \leqslant r \leqslant m.$$

Applying them to e_r, $m + 1 \leqslant r \leqslant 2m$, gives

$$\sum_{p=1}^{m} \alpha_p = \sum_{q=1}^{2m+1-r} \beta_q + \gamma \quad \text{for } m + 1 \leqslant r \leqslant 2m.$$

Finally, applying them to e_1 and e_{2m+1} gives

$$\sum_{p=1}^{m} \alpha_p = \sum_{q=1}^{m} \beta_q = 0.$$

It follows that $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \beta_1 = \beta_2 = \cdots = \beta_m = \gamma = 0$.

Case 2. $1 \leqslant t \leqslant n - 1$.

Sub-case 2(i). $n - t$ even, say $n - t = 2k$. We show that

$$\{A^{t+1}B, A^{t+2}B^2, \ldots, A^{t+k}B^k\} \cup \{BA^{t+1}, B^2A^{t+2}, \ldots, B^kA^{t+k}\}$$

is linearly independent (and so forms a basis for the subspace of matrices having zeros off the t-diagonal). Suppose that

$$\sum_{p=1}^{k} \alpha_p A^{t+p} B^p = \sum_{q=1}^{k} \beta_q B^q A^{t+q}.$$

Applying both sides to e_r, $2 \leqslant r \leqslant k$, to e_r, $k + 1 \leqslant r \leqslant 2k - 1$, to e_1 and to e_{2k} gives

$$\sum_{p=1}^{r-1} \alpha_p = \sum_{q=1}^{k} \beta_q \quad \text{for } 2 \leqslant r \leqslant k,$$

$$\sum_{p=1}^{k} \alpha_p = \sum_{q=1}^{2k-r} \beta_q \quad \text{for } k + 1 \leqslant r \leqslant 2k - 1,$$

$$\sum_{q=1}^{k} \beta_q = 0 \quad \text{and} \quad \sum_{p=1}^{k} \alpha_p = 0, \text{ respectively.}$$

It follows that $\alpha_1 = \alpha_2 = \cdots = \alpha_k = \beta_1 = \beta_2 = \cdots = \beta_k = 0$.

Sub-case 2(ii). $n - t$ odd, say $n - t = 2k + 1$. If $k = 0$, that is, $t = n - 1$, $\{A^{n-1}\}$ is linearly independent. Suppose that $k \geqslant 1$. We show that

$$\{A^t\} \cup \{A^{t+1}B, A^{t+2}B^2, \ldots, A^{t+k}B^k\} \cup \{BA^{t+1}, B^2A^{t+2}, \ldots, B^kA^{t+k}\}$$

is linearly independent. Suppose that

$$\alpha_0 A^t + \sum_{p=1}^{k} \alpha_p A^{t+p} B^p = \sum_{q=1}^{k} \beta_q B^q A^{t+q}.$$

Applying both sides to e_r, $2 \leqslant r \leqslant k + 1$, to e_r, $k + 2 \leqslant r \leqslant 2k$, to e_1 and to e_{2k+1} gives

$$\alpha_0 + \sum_{p=1}^{r-1} \alpha_p = \sum_{q=1}^{k} \beta_q \quad \text{for } 2 \leqslant r \leqslant k + 1,$$

$$\alpha_0 + \sum_{p=1}^{k} \alpha_p = \sum_{q=1}^{2k+1-r} \beta_q \quad \text{for } k + 2 \leqslant r \leqslant 2k,$$

$$\alpha_0 = \sum_{q=1}^{k} \beta_q \quad \text{and} \quad \alpha_0 + \sum_{p=1}^{k} \alpha_p = 0, \text{ respectively.}$$

It follows that $\alpha_0 = \alpha_1 = \cdots = \alpha_k = \beta_1 = \beta_2 = \cdots = \beta_k = 0$.

Case 3. $-(n - 1) \leqslant t \leqslant -1$. The desired result follows by symmetry. More precisely, let $\mathscr{E}_t$ be a set of words in A, B, each of length at most n, which forms a basis for the subspace of matrices having zeros off the $-t$-diagonal. Then $(\mathscr{E}_t)^* = \{E^* : E \in \mathscr{E}_t\}$ is a set of words in A, B, each of length at most n, which forms a basis for the subspace of matrices having zeros off the t-diagonal.

This completes the proof. $\square$

Remark. We cannot include the case $n = 3$ in the statement of the preceding proposition. It is easily checked that, when $n = 3$ and with A and B as in the statement of the proposition, the span of the words in A and B of length at most 3 is the subspace consisting of those matrices $C = (c_{i,j})$ satisfying $c_{2,2} = c_{1,1} + c_{3,3}$. Of course $\mathrm{msl}(A, B) = 4$ by Theorem 1 (or Theorem 2).

Proposition 4. *Let $n \geqslant 2$, $n \neq 3$, let B be the $n \times n$ strictly upper triangular elementary Jordan matrix, and let $A = B^*$. Then $M_n(\mathbb{C})$ is not spanned by words in A and B of length at most $n - 1$. Consequently* $\mathrm{msl}(A, B) = n$.

Proof. The result is obviously true if $n = 2$. Let $n \geqslant 4$. Every word in A and B maps each basis vector e_j to a basis vector or to the zero vector. Let W be a word in A and B of length at most $n-1$. We show that $(We_n|e_2) = (We_{n-1}|e_1)$. This is certainly true if both $(We_n|e_2)$ and $(We_{n-1}|e_1)$ are zero.

Suppose that $(We_n|e_2) \neq 0$. Then, since We_n is a basis vector, $We_n = e_2$. But any word with a factor of A in it would have to have length at least n to map e_n to e_2 (B shifts backwards and A shifts forwards). It follows that $W = B^{n-2}$. Then $(We_n|e_2) = (We_{n-1}|e_1) = 1$.

Suppose that $(We_{n-1}|e_1) \neq 0$. Then $We_{n-1} = e_1$. But any word with a factor of A in it would have to have length at least n to map e_{n-1} to e_1. It follows that $W = B^{n-2}$. Then $(We_n|e_2) = (We_{n-1}|e_1) = 1$.

Since $(We_n|e_2) = (We_{n-1}|e_1)$ for every word W in A and B of length at most $n-1$, the same equality holds for every element in the span of such words. Hence the span of such words cannot be $M_n(\mathbb{C})$. Thus $\mathrm{msl}(A, B) \geqslant n$ and it follows from Proposition 3 that $\mathrm{msl}(A, B) = n$. $\square$

The following case where the minimum spanning length of $\{A, B\}$ is also determinable seems worth mentioning.

Proposition 5. *Let* $\dim \mathscr{H} = n \geqslant 2$ *and let* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible and with* $\{A, B, AB, BA\}$ *linearly dependent. If* s_k *denotes the dimension of the span of the words in A and B of length at most* k, *then*

$$s_k - s_{k-1} = \begin{cases} k+1 & \text{if } 1 \leqslant k \leqslant n-1, \\ n & \text{if } k = n, \\ 2n-1-k & \text{if } n+1 \leqslant k \leqslant 2n-2, \end{cases}$$

taking $s_0 = 0$. *Consequently* $\mathrm{msl}(A, B) = 2n - 2$.

Proof. We can suppose that $\mathscr{H} = \mathbb{C}^n$. First we show that at least one of A, B is invertible. Suppose that B is not invertible. By Lemma 1, $BA \in \langle A, B, AB\rangle$. Let f be an eigenvector of B. Then $BAf \in \langle Af, Bf, ABf\rangle \subseteq \langle f, Af\rangle$ and, by induction, $BA^k f \in \langle f, Af, A^2 f, \ldots, A^k f\rangle$ for every $k \geqslant 0$. Thus the subspace $\langle f, Af, A^2 f, \ldots, A^{n-1} f\rangle$ is invariant under both A and B, so $\langle f, Af, A^2 f, \ldots, A^{n-1} f\rangle = \mathbb{C}^n$. Let $Be = 0$ with $e \neq 0$. Then $BAe \in \langle Ae, Be, ABe\rangle = \langle Ae\rangle$, so taking $f = Ae$ in the above we get that $\langle Ae, A^2 e, , A^3 e, \ldots, A^n e\rangle = \mathbb{C}^n$. Thus $\mathscr{R}(A) = \mathbb{C}^n$ and A is invertible.

We can assume that A is invertible. Let $BA = \alpha AB + \beta A + \gamma B$ with $\alpha, \beta, \gamma \in \mathbb{C}$. For every $k \geqslant 1$, let $\mathscr{W}_k$ denote the span of the words in A and B of length at most k and let

$$\mathscr{E}_k = \{A^p B^q : p, q \geqslant 0 \text{ and } 1 \leqslant p + q \leqslant k\}.$$

Clearly $\mathscr{W}_1 = \langle \mathscr{E}_1\rangle$. Assume that $\mathscr{W}_k = \langle \mathscr{E}_k\rangle$. If W is a word in A and B of length $k+1$, successively replacing a factor of BA by $\alpha AB + \beta A + \gamma B$ whenever possible, beginning with W, shows that W can be written as $W = \rho E + w$, where $\rho \in$

$\mathbb{C}$, $E \in \mathscr{E}_{k+1}$ and $w \in \mathscr{W}_k$. Since $\mathscr{W}_k = \langle \mathscr{E}_k \rangle \subseteq \langle \mathscr{E}_{k+1} \rangle$ it follows that $W \in \langle \mathscr{E}_{k+1} \rangle$. Hence, by induction, $\mathscr{W}_k = \langle \mathscr{E}_k \rangle$ for every $k \geqslant 1$.

Next we show that, if $k \geqslant n$,

$$\mathscr{W}_k = \big\langle \mathscr{E}_{n-1} \cup \{A^p B^q : p, q \leqslant n-1 \text{ and } n \leqslant p+q \leqslant k\} \cup \{I\} \big\rangle.$$

Let $k \geqslant n$ and let $\mathscr{V}_k = \langle \mathscr{E}_{n-1} \cup \{A^p B^q : p, q \leqslant n-1 \text{ and } n \leqslant p+q \leqslant k\} \cup \{I\} \rangle$. Since A is invertible, I can be written as a polynomial in A, of degree n, with no constant term. Thus $I \in \mathscr{W}_k$, so $\mathscr{V}_k \subseteq \mathscr{W}_k$.

Consider the word $A^p B^q$ where $p, q \geqslant 0$ and $n \leqslant p+q \leqslant k$. If $p \geqslant n$, A^p can be written as a polynomial in A of degree at most $n-1$, and if $q \geqslant n$, B^q can be written as a polynomial in B of degree at most $n-1$. It follows that $A^p B^q$ is linear combination of words of the form $A^{p'} B^{q'}$ where $0 \leqslant p', q' \leqslant n-1$ and $p'+q' \leqslant p+q$. Hence $A^p B^q \in \mathscr{V}_k$ and so $\mathscr{V}_k = \mathscr{W}_k$.

Using the fact that I can be written as a polynomial in A, of degree n, with no constant term it follows that, if $k \geqslant n$,

$$\mathscr{W}_k = \big\langle \mathscr{E}_{n-1} \cup \{A^p B^q : p, q \leqslant n-1 \text{ and } n \leqslant p+q \leqslant k\} \cup \{A^n\} \big\rangle.$$

From this it in turn follows that, for $k \geqslant 2n-2$,

$$\mathscr{W}_k = \mathscr{W}_{2n-2} = \big\langle \{A^p B^q : 0 \leqslant p, q \leqslant n-1 \text{ and } p+q \geqslant 1\} \cup \{A^n\} \big\rangle.$$

Thus $\mathscr{W}_{2n-2} = M_n(\mathbb{C})$ and $\{A^p B^q : 0 \leqslant p, q \leqslant n-1 \text{ and } p+q \geqslant 1\} \cup \{A^n\}$ is a basis for $M_n(\mathbb{C})$ (since it spans $M_n(\mathbb{C})$ and has n^2 elements). It follows also that, if $1 \leqslant k \leqslant n-1$, then $\mathscr{E}_k$ is a basis for $\mathscr{W}_k$ and if $k \geqslant n$ then $\{A^p B^q : p, q \leqslant n-1 \text{ and } n \leqslant p+q \leqslant k\} \cup \{A^n\}$ is a basis for $\mathscr{W}_k$. It readily follows that $s_k - s_{k-1}$ satisfies the condition stated in the proposition. Clearly $\operatorname{msl}(A, B) \leqslant 2n-2$. Since $s_{2n-2} - s_{2n-3} = 1$, we have $\operatorname{msl}(A, B) = 2n-2$. $\square$

The following example shows that on every space $\mathscr{H}$ with $\dim \mathscr{H} = n \geqslant 2$ there exist $A, B \in \mathscr{B}(\mathscr{H})$ with $\{A, B\}$ irreducible such that $\{A, B, AB, BA\}$ is linearly dependent (and neither A nor B is unicellular).

Example 3. Let $n \geqslant 2$ and put $\alpha = \mathrm{e}^{2\pi \mathrm{i}/n}$ and $\lambda_k = 1 + \alpha + \alpha^2 + \cdots + \alpha^k$, $1 \leqslant k \leqslant n-1$. Let

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

and let $B = \operatorname{diag}(0, 1, \lambda_1, \lambda_2, \ldots, \lambda_{n-2})$. Using the facts that $1 + \alpha\lambda_k = \lambda_{k+1}$, $1 \leqslant k \leqslant n-2$ and $\lambda_{n-1} = 0$ it is easily checked that $BA = \alpha AB + A$.

The nontrivial invariant subspaces of B are the subspaces $\langle\{e_i : i \in \mathscr{S}\}\rangle$ where $\mathscr{S}$ is a nontrivial subset of $\{1, 2, \ldots, n\}$. Each basis vector e_i is a cyclic vector for A so $\{A, B\}$ is irreducible. By Proposition 5, $\mathrm{msl}(A, B) \leqslant 2n - 2$. Since neither A nor B is unicellular, this fact does not follow from Theorem 2. (The characteristic polynomial of A is $\lambda^n - 1 = 0$, so A has a basis consisting of eigenvectors.)

In Proposition 2 we showed that, if $\{A, B\}$ is an irreducible pair of transformations on a space of dimension 2, the dimension of the span of the words in A, B of length at most 2 is 4. It may be of interest to know what the dimension of this span is on higher dimensional spaces. This is answered by the following proposition and example.

Proposition 6. *Let* $\dim \mathscr{H} \geqslant 3$ *and let* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible. The dimension of the span of the words in A and B of length at most* 2 *is* 5 *or* 6.

Proof. By Lemma 1, $\{A, B, AB\}$ is linearly independent. Let $Ae = \lambda e$ and $Bf = \mu f$ with $e, f \neq 0$ and μ, λ scalars.

If $BA \in \langle A, B, AB\rangle$ then the dimension of $\langle A, B, AB, BA, A^2, B^2\rangle$ is 5 by Proposition 5.

Suppose that $BA \notin \langle A, B, AB\rangle$. Assume that A^2 and B^2 both belong to $\langle A, B, AB, BA\rangle$. Then $A^2 = \alpha A + \beta B + \gamma AB + \delta BA$ and $B^2 = \alpha' A + \beta' B + \gamma' AB + \delta' BA$ for some scalars $\alpha, \beta, \ldots, \delta'$. Then $\lambda^2 e = \alpha\lambda e + \beta Be + \gamma ABe + \delta\lambda Be$ and $B^2 e = \alpha'\lambda e + \beta' Be + \gamma' ABe + \delta'\lambda Be$. If $\gamma \neq 0$ then $\langle e, Be\rangle$ would be invariant under A and B. Hence $\gamma = 0$ so $A^2 \in \langle A, B, BA\rangle$. By Lemma 1 it follows that A^2 is a scalar multiple of A. Similarly B^2 is a scalar multiple of B. But, since $\dim \mathscr{H} \geqslant 3$, every pair of quadratic transformations on $\mathscr{H}$ has a common nontrivial invariant subspace by [1, Theorem 2]. This contradiction shows that either $A^2 \notin \langle A, B, AB, BA\rangle$ or $B^2 \notin \langle A, B, AB, BA\rangle$ so the dimension of $\langle A, B, AB, BA, A^2, B^2\rangle$ is 5 or 6. □

Corollary 2. *If* $\dim \mathscr{H} = n \geqslant 3$ *and* $A, B \in \mathscr{B}(\mathscr{H})$ *with* $\{A, B\}$ *irreducible then* $\mathrm{msl}(A, B) \leqslant n^2 - 3$.

Proof. With notation as in the proof of Proposition 1, $a = n^2$, and we have $s_3 \geqslant s_2 + 1, s_4 \geqslant s_2 + 2, \ldots, n^2 = s_{p_0} \geqslant s_2 + p_0 - 2$. But $s_2 \geqslant 5$ by Proposition 6, and $p_0 = \mathrm{msl}(A, B)$ so $n^2 \geqslant 3 + \mathrm{msl}(A, B)$. □

Example 3 shows that, if $\{A, B\}$ is an irreducible pair of transformations on a space of dimension at least 3, the dimension of the span of the words in A, B of length at most 2 can be 5. The following example shows that it can also be 6.

Example 4. Let $n \geqslant 3$ and let B be the $n \times n$ strictly upper triangular elementary Jordan matrix. Let $A = B^*$. The nontrivial invariant subspaces of B are $\langle e_1\rangle$, $\langle e_1, e_2\rangle$, $\langle e_1, e_2, e_3\rangle, \ldots, \langle e_1, e_2, \ldots, e_{n-1}\rangle$. Clearly $\{A, B\}$ is irreducible. The 3, 1-entries of A, B, AB, A^2 are 0, 0, 0, 1, respectively, so $A^2 \notin \langle A, B, AB\rangle$. The 1, 1-

entries of A, B, AB, A^2, BA are 0, 0, 0, 0, 1, respectively, so $BA \notin \langle A, B, AB, A^2\rangle$. The 1, 3-entries of A, B, AB, A^2, BA, B^2 are 0, 0, 0, 0, 0, 1, respectively, so $B^2 \notin \langle A, B, AB, A^2, BA\rangle$. Using Lemma 1, it follows that the dimension of $\langle A, B, AB, A^2, BA, B^2\rangle$ is 6.

Remark. None of the above examples or results leads to an example of an irreducible pair $\{A, B\}$ of $n \times n$ matrices (with $n \geqslant 2$) with minimum spanning length greater than $2n - 2$. As far as the author is aware, no such example has yet been found. This supports the conjecture that no such pair exists, that is, that $\mathrm{msl}(A, B) \leqslant 2n - 2$ for every irreducible pair $\{A, B\}$.

Acknowledgements

The author would like to thank H. Radjavi and P. Rosenthal for their encouragement and useful discussions regarding the preparation of this paper.

References

[1] L. Brickman, P.A. Fillmore, The invariant subspace lattice of a linear transformation, Canad. J. Math. 19 (1967) 810–822.
[2] E.A. Nordgren, M. Radjabalipour, H. Radjavi, P. Rosenthal, Quadratic operators and invariant subspaces, Studia Math. 88 (1988) 263–268.
[3] H. Radjavi, P. Rosenthal, Simultaneous Triangularization, Springer-Verlag, New York, 2000.